



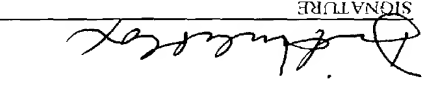
101-2

09744313.082401

526 Rec'd PCT/PTO 18 JAN 2001

U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE <b>TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371</b>		ATTORNEY'S DOCKET NUMBER <b>PF-0706 USN</b>
INTERNATIONAL APPLICATION NO. PCT/US00/14831		U.S. APPLICATION NO. (known, see 37 CFR 1.55) TO BE ASSIGNED <b>09/744313</b>
INTERNATIONAL FILING DATE 26 May 2000		PRIORITY DATE CLAIMED 27 May 1999
TITLE OF INVENTION <b>SORTING NEXINS</b>		
APPLICANT(S) FOR DO/EO/US <b>INCYTE GENOMICS, INC.; YUE, Henry; TANG, Y. Tom; AZIMZAI, Yalda</b>		
<p>Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:</p> <ol style="list-style-type: none"><li>1. <input checked="" type="checkbox"/> This is the <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. 371.</li><li>2. <input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. 371.</li><li>3. <input type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).</li><li>4. <input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).</li><li>5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))<ol style="list-style-type: none"><li>a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau)</li><li>b. <input type="checkbox"/> has been communicated by the International Bureau.</li><li>c. <input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</li></ol></li><li>6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).</li><li>7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))<ol style="list-style-type: none"><li>a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).</li><li>b. <input type="checkbox"/> have been communicated by the International Bureau.</li><li>c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</li><li>d. <input checked="" type="checkbox"/> have not been made and will not be made.</li></ol></li><li>8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).</li><li>9. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).</li><li>10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).</li></ol>		
<b>Items 11 to 16 below concern document(s) or information included:</b>		
<ol style="list-style-type: none"><li>11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</li><li>12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.</li><li>13. <input type="checkbox"/> A <b>FIRST</b> preliminary amendment. <input type="checkbox"/> A <b>SECOND</b> or <b>SUBSEQUENT</b> preliminary amendment.</li><li>14. <input type="checkbox"/> A substitute specification.</li><li>15. <input type="checkbox"/> A change of power of attorney and/or address letter.</li><li>16. <input checked="" type="checkbox"/> Other items or information:  1) Transmittal Letter (2 pp, in duplicate) 2) Return Postcard 3) Express Mail Label No.: <b>EL 743 380 058 US</b></li></ol>		

JC07 Rec'd PCT/PTO 18 JAN 2001

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)		INTERNATIONAL APPLICATION NO.		ATTORNEY'S DOCKET NUMBER	
09/744313		PCT/US00/14831		PF-0706 USN	
<b>17. The following fees are submitted:</b>					
<b>BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):</b> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO USPTO but International Search Report not prepared by the EPO or JPO.....\$1000.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO.....\$860.00 International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4).....\$690.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....\$100.00					
<b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b>					
\$860.00					
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(c)).					
CLAIMS					
NUMBER FILED		NUMBER EXTRA		RATE	
33		13		\$234.00	
Total Claims		=		X \$ 18.00	
Independent Claims		=		X \$ 80.00	
0				\$	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)					
				+ \$270.00	
				\$	
<b>TOTAL OF ABOVE CALCULATIONS =</b>					
\$1094.00					
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.					
<b>SUBTOTAL =</b>					
\$1094.00					
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).					
TOTAL NATIONAL FEE =					
\$1094.00					
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property +					
TOTAL FEES ENCLOSED =					
\$1094.00					
Amount to be Refunded:		\$			
Charged:		\$			
a. <input type="checkbox"/> A check in the amount of \$_____ to cover the above fees is enclosed. b. <input type="checkbox"/> Please charge my Deposit Account No. 09-0108 in the amount of \$ 1094.00 to cover the above fees. c. <input type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed.					
NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.					
SEND ALL CORRESPONDENCE TO:					
INCYTE GENOMICS, INC. 3160 Porter Drive Palo Alto, CA 94304					
NAME: Diana Hamlet-Cox					
SIGNATURE: 					
REGISTRATION NUMBER: 33,302					
DATE: 16 January 2001					

0442300 21 2 4 5 6 7 8 9 0

WO 00/73334

PCT/US00/14831

## SORTING NEXINS

09/744313

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of sorting nexins and to the use of these sequences in the diagnosis, treatment, and prevention of immune, neurological, gastrointestinal, genetic, and smooth muscle disorders, and cell proliferative disorders, including cancer.

### BACKGROUND OF THE INVENTION

The organization of a cell is designed for optimal structure and function. Each molecule produced by the cell must be targeted to the appropriate subcellular location. To be properly localized, each molecule must contain information which provides its cellular address. Trafficking machinery, which includes the sorting nexin protein family, can use this information to appropriately transport the molecule to its destination. This is a dynamic process, and the location of a molecule can change over time. Defects in any step involving a molecule's proper localization can result in a disorder, such as diabetes or Alzheimer's disease (James, D.E. and R.C. Piper (1994) J. Cell Biol. 126:1123-1126; Nordstedt, C. et al. (1993) J. Biol. Chem. 268:608-612).

The information that provides the address for targeting resides within the primary structure of the protein. Certain amino acid sequences act as "delivery codes" during the processing or recycling of a molecule and assure that the molecule is properly localized. Several motifs have been identified. The two principal motifs are the nuclear localization signal and signal peptide. Nuclear localization signals (NLSs) consist of short stretches of amino acids enriched in basic residues. NLSs are found on proteins that are targeted to the nucleus, such as the glucocorticoid receptor. The NLS is recognized by the NLS receptor, importin, which then interacts with the monomeric GTP-binding protein Ran. This NLS protein/receptor/Ran complex navigates the nuclear pore with the help of the homodimeric protein nuclear transport factor 2 (NTF2). NTF2 binds to the GDP-bound form of Ran and to multiple proteins of the nuclear pore complex, such as p62.

Signal peptides are found on proteins that are targeted to the endoplasmic reticulum (ER). Signal peptides consist of stretches of amino acids enriched in hydrophobic residues. Signal peptides are usually found at the extreme N-terminus of the protein and are recognized by a cytosolic signal-recognition peptide (SRP). The SRP binds to the signal peptide and to an SRP receptor, an integral membrane protein in the ER. Once bound to the SRP receptor, the newly formed protein containing the signal peptide is translocated across the ER membrane. Proteins containing signal peptides may eventually be inserted into the lipid bilayer, translocated to the lumen of an organelle, or secreted from the cell.







WO 00/73334

PCT/US00/14831

WW domain is composed of beta strands grouped around four conserved aromatic residues, generally tryptophan.

Like SH3, the SH2 domain is defined by homology to a region of c-Src. SH2 domains interact directly with phospho-tyrosine residues, thus providing an immediate mechanism for the regulation and transduction of receptor tyrosine kinase-mediated signaling pathways. For example, as many as ten distinct SH2 domains are capable of binding to phosphorylated tyrosine residues in the activated PDGF receptor, thereby providing a highly coordinated and finely tuned response to ligand-mediated receptor activation. (Reviewed in Schaffhausen, B. (1995) *Biochim. Biophys. Acta.* 1242:61-75.)

The pleckstrin homology domain (PHD) was originally identified in pleckstrin, the predominant substrate for protein kinase C in platelets. Since its discovery, this domain has been identified in over 90 proteins involved in intracellular signaling or cytoskeletal organization. Proteins containing the pleckstrin homology domain include a variety of kinases, phospholipase-C isoforms, guanine nucleotide release factors, and GTPase activating proteins. Many PHD proteins function in association with the plasma membrane, and this association appears to be mediated by the PHD itself. PHDs share a common structure composed of two antiparallel beta sheets flanked by an amphipathic alpha helix. Variable loops connecting the component beta strands generally occur within a positively charged environment and may function as ligand binding sites (Lemmon, M.A. et al. (1996) *Cell* 85:621-624).

Ankyrin (ANK) repeats mediate protein-protein interactions associated with diverse intracellular signaling functions. For example, ANK repeats are found in proteins involved in cell proliferation such as kinases, kinase inhibitors, tumor suppressors, and cell cycle control proteins. (See, for example, Kalus, W. et al. (1997) *FEBS Lett.* 401:127-132; Ferrante, A.W. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:1911-1915.) These proteins generally contain multiple ANK repeats, each composed of about 33 amino acids. Myotrophin is an ANK repeat protein that plays a key role in the development of cardiac hypertrophy, a contributing factor to many heart diseases. Structural studies show that the myotrophin ANK repeats, like other ANK repeats, each form a helix-turn-helix core preceded by a protruding "tip." These tips are of variable sequence and may play a role in protein-protein interactions. The helix-turn-helix region of the ANK repeats stack on top of one another and are stabilized by hydrophobic interactions (Yang, Y. et al. (1998) *Structure* 6:619-626).

The discovery of new sorting nexins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of immune, neurological, gastrointestinal, genetic, and smooth muscle disorders, and cell proliferative disorders, including cancer.

35

**SUMMARY OF THE INVENTION**

The invention features purified polypeptides, sorting nexins, referred to collectively as "SNEXN" and individually as "SNEXN-1" and "SNEXN-2." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence of SEQ ID NO:1, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence of SEQ ID NO:1, c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1.

10 The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence of SEQ ID NO:1, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence of SEQ ID NO:1, c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1. In one alternative, the polynucleotide encodes a polypeptide of SEQ ID NO:1. In another alternative, the polynucleotide comprises a polynucleotide sequence of SEQ ID NO:3.

15 Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence of SEQ ID NO:1, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence of SEQ ID NO:1, c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

25 The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence of SEQ ID NO:1, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence of SEQ ID NO:1, c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

30 Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence of SEQ ID NO:1, b) a naturally occurring amino acid sequence having at least 90%

WO 00/73334

PCT/US00/14831

sequence identity to an amino acid sequence of SEQ ID NO:1, c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1.

The invention further provides an isolated polynucleotide comprising a polynucleotide  
 5 sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60  
 10 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide sequence having at least 70%  
 15 sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide,  
 20 under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample,  
 25 said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target  
 30 polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a pharmaceutical composition comprising an effective amount  
 35 of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an



of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

5 The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence of SEQ ID NO:1, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence of SEQ ID NO:1, c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

15 The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:3-4, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

20 The invention further provides an isolated polynucleotide comprising a polynucleotide sequence of SEQ ID NO:4. Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide comprising a polynucleotide sequence of SEQ ID NO:4. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

25 The invention also provides a method for producing a polypeptide comprising an amino acid sequence of SEQ ID NO:2. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide comprising a polynucleotide sequence of SEQ ID NO:4, and b) recovering the polypeptide so expressed.

## BRIEF DESCRIPTION OF THE TABLES

35 Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-

WO 00/73334

PCT/US00/14831

length sequences encoding SNEXN.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of SNEXN.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific  
5 expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding SNEXN were isolated.

10 Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

### DESCRIPTION OF THE INVENTION

15 Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

20 It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

25 Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing  
30 the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

### DEFINITIONS

35 "SNEXN" refers to the amino acid sequences of substantially purified SNEXN obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and











WO 00/73334

PCT/US00/14831

example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:3-4 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:3-4 and the region of SEQ ID NO:3-4 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

5 A fragment of SEQ ID NO:1-2 is encoded by a fragment of SEQ ID NO:3-4. A fragment of SEQ ID NO:1-2 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-2. For example, a fragment of SEQ ID NO:1-2 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-2. The precise length of a fragment of SEQ ID NO:1-2 and the region of SEQ ID NO:1-2 to which the fragment corresponds are  
10 routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

15 "Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps  
20 in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of  
25 molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent  
30 similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at  
35 <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence



WO 00/73334

PCT/US00/14831

sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by

5 CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

10 *Matrix: BLOSUM62*  
*Open Gap: 11 and Extension Gap: 1 penalties*  
*Gap x drop-off: 50*  
*Expect: 10*  
*Word Size: 3*  
15 *Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least  
20 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for  
25 chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a  
30 complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific  
35 binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive

WO 00/73334

PCT/US00/14831

conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 5 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of 10 the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present 15 invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic 20 solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

25 The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g.,  $C_0t$  or  $R_0t$  analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate 30 to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression 35 of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect

WO 00/73334

PCT/US00/14831

cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of SNEXN which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of SNEXN which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of SNEXN. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of SNEXN.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an SNEXN may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of SNEXN.

"Probe" refers to nucleic acid sequences encoding SNEXN, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.



WO 00/73334

PCT/US00/14831

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

5 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the  
10 specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al., 1990, PCR  
15 Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such  
20 purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of  
25 Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the  
30 selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved  
35 regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both





WO 00/73334

PCT/US00/14831

introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

5 A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or  
10 greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the  
15 reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The  
20 presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-  
25 1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

## THE INVENTION

The invention is based on the discovery of new human sorting nexins (SNEXN), the  
30 polynucleotides encoding SNEXN, and the use of these compositions for the diagnosis, treatment, or prevention of immune, neurological, gastrointestinal, genetic, and smooth muscle disorders, and cell proliferative disorders, including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding SNEXN. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the  
35 polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte

WO 00/73334

PCT/US00/14831

clones in which nucleic acids encoding each SNEXN were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus  
 5 nucleotide sequence of each SNEXN and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6  
 10 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions  
 15 associated with nucleotide sequences encoding SNEXN. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:3-4 and to distinguish between SEQ ID NO:3-4 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column  
 20 3 lists tissue categories which express SNEXN as a fraction of total tissues expressing SNEXN. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing SNEXN as a fraction of total tissues expressing SNEXN. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries  
 25 from which cDNA clones encoding SNEXN were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:3 maps to chromosome 6 within the interval from 91.8 to 96.1 centiMorgans. This interval also contains genes and ESTs associated with maple syrup urine disease, caused by a  
 30 defect in the E1-beta subunit of branched chain alpha-keto acid dehydrogenase. SEQ ID NO:4 maps to chromosome 3 within the interval from 140.00 to 148.70 centiMorgans.

The invention also encompasses SNEXN variants. A preferred SNEXN variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the SNEXN amino acid sequence, and which contains at least one functional or  
 35 structural characteristic of SNEXN.

WO 00/73334

PCT/US00/14831

The invention also encompasses polynucleotides which encode SNEXN. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4, which encodes SNEXN. The polynucleotide sequences of SEQ ID NO:3-4, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SNEXN. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SNEXN. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:3-4. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SNEXN.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding SNEXN, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SNEXN, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SNEXN and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SNEXN under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SNEXN or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SNEXN and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode SNEXN and SNEXN derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the

WO 00/73334

PCT/US00/14831

synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding SNEXN or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of  
 5 hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:3-4 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found  
 15 in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing  
 20 system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding SNEXN may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.)

Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme  
 35







28

WO 00/73334

PCT/US00/14831

adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al., (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.)  
The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SNEXN. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SNEXN can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding SNEXN into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of SNEXN are needed, e.g. for the production of antibodies, vectors which direct high level expression of SNEXN may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of SNEXN. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, *supra*; and Scorer, *supra*.)

Plant systems may also be used for expression of SNEXN. Transcription of sequences encoding SNEXN may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, *supra*; Broglie, *supra*; and Winter, *supra*.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SNEXN may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader

WO 00/73334

PCT/US00/14831

sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses SNEXN in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of SNEXN in cell lines is preferred. For example, sequences encoding SNEXN can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is



WO 00/73334

PCT/US00/14831

produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode SNEXN may be designed to contain signal sequences which direct secretion of SNEXN through a prokaryotic or eukaryotic cell membrane.

5 In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity.

10 Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding SNEXN may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric SNEXN protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SNEXN activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SNEXN encoding sequence and the heterologous protein sequence, so that SNEXN may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled SNEXN may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid

WO 00/73334

PCT/US00/14831

precursor, for example, <sup>35</sup>S-methionine.

SNEXN of the present invention or fragments thereof may be used to screen for compounds that specifically bind to SNEXN. At least one and up to a plurality of test compounds may be screened for specific binding to SNEXN. Examples of test compounds include antibodies,  
5 oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of SNEXN, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which SNEXN  
10 binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express SNEXN, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing SNEXN or cell membrane fractions which contain SNEXN are then  
15 contacted with a test compound and binding, stimulation, or inhibition of activity of either SNEXN or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with SNEXN, either in  
20 solution or affixed to a solid support, and detecting the binding of SNEXN to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

SNEXN of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of SNEXN. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SNEXN activity, wherein SNEXN is combined with at least one test compound, and the activity of SNEXN in the presence of a test compound is compared with the activity of SNEXN in the absence  
30 of the test compound. A change in the activity of SNEXN in the presence of the test compound is indicative of a compound that modulates the activity of SNEXN. Alternatively, a test compound is combined with an in vitro or cell-free system comprising SNEXN under conditions suitable for SNEXN activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SNEXN may do so indirectly and need not come in direct contact with the  
35 test compound. At least one and up to a plurality of test compounds may be screened.





WO 00/73334

PCT/US00/14831

genetic, and smooth muscle disorders, and cell proliferative disorders, including cancer. In the treatment of disorders associated with increased SNEXN expression or activity, it is desirable to decrease the expression or activity of SNEXN. In the treatment of disorders associated with decreased SNEXN expression or activity, it is desirable to increase the expression or activity of

5 SNEXN.

Therefore, in one embodiment, SNEXN or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SNEXN. Examples of such disorders include, but are not limited to, an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's

10 disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's

15 thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic

20 lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease

25 and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-

30 Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders,

35 peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic,

WO 00/73334

PCT/US00/14831

endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a gastrointestinal disorder such as esophagitis, esophageal carcinoma, gastritis, gastric carcinoma, inflammatory bowel disease, cholecystitis, infections of the intestinal tract, pancreatitis, pancreatic carcinoma, cirrhosis, hepatitis, hepatoma, colitis, colonic carcinoma, and Crohn's disease; a genetic disorder such as adrenoleukodystrophy, Alport's syndrome, choroideremia, Duchenne and Becker muscular dystrophy, Down's syndrome, cystic fibrosis, chronic granulomatous disease, Gaucher's disease, Huntington's chorea, Marfan's syndrome, muscular dystrophy, myotonic dystrophy, pycnodysostosis, Refsum's syndrome, retinoblastoma, sickle cell anemia, thalassemia, Werner syndrome, von Willebrand's disease, Wilms' tumor, Zellweger syndrome, peroxisomal acyl-CoA oxidase deficiency, peroxisomal thiolase deficiency, peroxisomal bifunctional protein deficiency, mitochondrial carnitine palmitoyl transferase and carnitine deficiency, mitochondrial very-long-chain acyl-CoA dehydrogenase deficiency, mitochondrial medium-chain acyl-CoA dehydrogenase deficiency, mitochondrial short-chain acyl-CoA dehydrogenase deficiency, mitochondrial electron transport flavoprotein and electron transport flavoprotein:ubiquinone oxidoreductase deficiency, mitochondrial trifunctional protein deficiency, and mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency; a smooth muscle disorder such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia, with smooth muscle including, but not limited to, that of the blood vessels, gastrointestinal tract, heart, and uterus; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing SNEXN or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SNEXN including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified

WO 00/73334

PCT/US00/14831

SNEXN in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SNEXN including, but not limited to, those provided above.

5 In still another embodiment, an agonist which modulates the activity of SNEXN may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SNEXN including, but not limited to, those listed above.

10 In a further embodiment, an antagonist of SNEXN may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SNEXN. Examples of such disorders include, but are not limited to, those immune, neurological, gastrointestinal, genetic, and smooth muscle disorders, and cell proliferative disorders, including cancer, described above. In one aspect, an antibody which specifically binds SNEXN may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SNEXN.

15 In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SNEXN may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SNEXN including, but not limited to, those described above.

20 In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

25 An antagonist of SNEXN may be produced using methods which are generally known in the art. In particular, purified SNEXN may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SNEXN. Antibodies to SNEXN may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

30 For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with SNEXN or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic

WO 00/73334

PCT/US00/14831

polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to SNEXN have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of SNEXN amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to SNEXN may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SNEXN-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for SNEXN may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either



WO 00/73334

PCT/US00/14831

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 5 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other 10 systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding SNEXN may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency 15 (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene* 20 *Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and Somia, N. (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites 25 (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in SNEXN expression or regulation causes disease, the expression of 30 SNEXN from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SNEXN are treated by constructing mammalian expression vectors encoding SNEXN and introducing these vectors by mechanical means into SNEXN-deficient cells. Mechanical transfer technologies for 35 use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii)

WO 00/73334

PCT/US00/14831

ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

- 5 Expression vectors that may be effective for the expression of SNEXN include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SNEXN may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the
- 10 (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the
- 15 FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SNEXN from a normal individual.

- Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver
- 20 polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

- 25 In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to SNEXN expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SNEXN under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences
- 30 required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al.
- 35 (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and







WO 00/73334

PCT/US00/14831

engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding SNEXN.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding SNEXN. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2'O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SNEXN. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SNEXN expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SNEXN may be therapeutically useful, and in the treatment of disorders





WO 00/73334

PCT/US00/14831

routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example SNEXN or fragments thereof, antibodies of SNEXN, and agonists, antagonists or inhibitors of SNEXN, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be  
5 determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and  
10 animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the  
15 subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4  
20 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.  
25 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

#### DIAGNOSTICS

In another embodiment, antibodies which specifically bind SNEXN may be used for the  
30 diagnosis of disorders characterized by expression of SNEXN, or in assays to monitor patients being treated with SNEXN or agonists, antagonists, or inhibitors of SNEXN. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SNEXN include methods which utilize the antibody and a label to detect SNEXN in human body fluids or in extracts of cells or tissues. The antibodies may be used with or  
35 without modification, and may be labeled by covalent or non-covalent attachment of a reporter

WO 00/73334

PCT/US00/14831

molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring SNEXN, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SNEXN expression.

5 Normal or standard values for SNEXN expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to SNEXN under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SNEXN expressed in subject, control, and disease samples from biopsied tissues are compared with the  
10 standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding SNEXN may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect  
15 and quantify gene expression in biopsied tissues in which expression of SNEXN may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of SNEXN, and to monitor regulation of SNEXN levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SNEXN or closely related molecules may be used  
20 to identify nucleic acid sequences which encode SNEXN. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding SNEXN, allelic variants, or related sequences.

25 Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the SNEXN encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:3-4 or from genomic sequences including promoters, enhancers, and introns of the SNEXN gene.

Means for producing specific hybridization probes for DNAs encoding SNEXN include the  
30 cloning of polynucleotide sequences encoding SNEXN or SNEXN derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as <sup>32</sup>P or <sup>35</sup>S, or by enzymatic labels,  
35 such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

WO 00/73334

PCT/US00/14831

Polynucleotide sequences encoding SNEXN may be used for the diagnosis of disorders associated with expression of SNEXN. Examples of such disorders include, but are not limited to, an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a gastrointestinal disorder such as esophagitis, esophageal carcinoma, gastritis, gastric carcinoma, inflammatory bowel disease, cholecystitis, infections of the intestinal

WO 00/73334

PCT/US00/14831

tract, pancreatitis, pancreatic carcinoma, cirrhosis, hepatitis, hepatoma, colitis, colonic carcinoma, and Crohn's disease; a genetic disorder such as adrenoleukodystrophy, Alport's syndrome, choroideremia, Duchenne and Becker muscular dystrophy, Down's syndrome, cystic fibrosis, chronic granulomatous disease, Gaucher's disease, Huntington's chorea, Marfan's syndrome, muscular  
5 dystrophy, myotonic dystrophy, pycnodysostosis, Refsum's syndrome, retinoblastoma, sickle cell anemia, thalassemia, Werner syndrome, von Willebrand's disease, Wilms' tumor, Zellweger syndrome, peroxisomal acyl-CoA oxidase deficiency, peroxisomal thiolase deficiency, peroxisomal bifunctional protein deficiency, mitochondrial carnitine palmitoyl transferase and carnitine deficiency, mitochondrial very-long-chain acyl-CoA dehydrogenase deficiency, mitochondrial  
10 medium-chain acyl-CoA dehydrogenase deficiency, mitochondrial short-chain acyl-CoA dehydrogenase deficiency, mitochondrial electron transport flavoprotein and electron transport flavoprotein:ubiquinone oxidoreductase deficiency, mitochondrial trifunctional protein deficiency, and mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency; a smooth muscle disorder such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's  
15 syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia, with smooth muscle including, but not limited to, that of the blood vessels, gastrointestinal tract, heart, and uterus; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective  
20 tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen,  
25 testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding SNEXN may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SNEXN expression. Such qualitative or quantitative methods are well known in the art.

30 In a particular aspect, the nucleotide sequences encoding SNEXN may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SNEXN may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a  
35 standard value. If the amount of signal in the patient sample is significantly altered in comparison to



WO 00/73334

PCT/US00/14831

a control sample then the presence of altered levels of nucleotide sequences encoding SNEXN in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

5 In order to provide a basis for the diagnosis of a disorder associated with expression of SNEXN, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SNEXN, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from  
10 normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated,  
15 hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or  
20 overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

25 Additional diagnostic uses for oligonucleotides designed from the sequences encoding SNEXN may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding SNEXN, or a fragment of a polynucleotide complementary to the polynucleotide encoding SNEXN, and will be employed under optimized conditions for identification of a specific gene or  
30 condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding SNEXN may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic  
35 disease in humans. Methods of SNP detection include, but are not limited to, single-stranded

WO 00/73334

PCT/US00/14831

conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding SNEXN are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of SNEXN include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for SNEXN, or SNEXN or fragments thereof may

WO 00/73334

PCT/US00/14831

be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SNEXN may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SNEXN on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely

WO 00/73334

PCT/US00/14831

localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, SNEXN, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between SNEXN and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SNEXN, or fragments thereof, and washed. Bound SNEXN is then detected by methods well known in the art. Purified SNEXN can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SNEXN specifically compete with a test compound for binding SNEXN. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SNEXN.

In additional embodiments, the nucleotide sequences which encode SNEXN may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications mentioned above and below, in particular U.S. Ser. No. 60/136,740 and U.S. Ser. No. 60/139,566, are hereby expressly incorporated by reference.

## EXAMPLES





WO 00/73334

PCT/US00/14831

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:3-4. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is

WO 00/73334

PCT/US00/14831

calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding SNEXN occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### V. Chromosomal Mapping of SNEXN Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:3-4 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:3-4 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

The genetic map locations of SEQ ID NO:3 and SEQ ID NO:4 are described in The Invention as ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in



WO 00/73334

PCT/US00/14831

each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention where applicable.

#### VI. Extension of SNEXN Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:3-4 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%)

WO 00/73334

PCT/US00/14831

agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on  
 5 antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min;  
 10 Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM  
 15 BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, the polynucleotide sequences of SEQ ID NO:3-4 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

## VII. Labeling and Use of Individual Hybridization Probes

20 Hybridization probes derived from SEQ ID NO:3-4 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine  
 25 triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I,  
 30 Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.  
 35 Hybridization patterns are visualized using autoradiography or an alternative imaging means and

compared.

### VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the  
 5    aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding  
 10    procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof  
 15    may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a  
 20    fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

#### 25    Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40  
 30    μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and  
 35    incubated for 20 minutes at 85 °C to stop the reaction and degrade the RNA. Samples are purified

WO 00/73334

PCT/US00/14831

using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and

5 resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

#### Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are

10 amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water

15 washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US

20 Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water.

25 Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

#### Hybridization

Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and

30 Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is

35 incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash

WO 00/73334

PCT/US00/14831

buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

#### Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an  
5 Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines  
at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is  
focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide  
containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-  
scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a  
10 resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially.  
Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477,  
Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate  
filters positioned between the array and the photomultiplier tubes are used to filter the signals. The  
15 emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is  
typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source,  
although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a  
cDNA control species added to the sample mixture at a known concentration. A specific location on  
20 the array contains a complementary DNA sequence, allowing the intensity of the signal at that  
location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples  
from different sources (e.g., representing test and control cells), each labeled with a different  
fluorophore, are hybridized to a single array for the purpose of identifying genes that are  
differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the  
25 two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital  
(A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC  
computer. The digitized data are displayed as an image where the signal intensity is mapped using a  
linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high  
30 signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and  
measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping  
emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each  
spot is centered in each element of the grid. The fluorescence signal within each element is then  
35 integrated to obtain a numerical value corresponding to the average intensity of the signal. The

WO 00/73334

PCT/US00/14831

software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

#### IX. Complementary Polynucleotides

Sequences complementary to the SNEXN-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SNEXN. Although use of  
 5 oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SNEXN. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a  
 10 complementary oligonucleotide is designed to prevent ribosomal binding to the SNEXN-encoding transcript.

#### X. Expression of SNEXN

Expression and purification of SNEXN is achieved using bacterial or virus-based expression systems. For expression of SNEXN in bacteria, cDNA is subcloned into an appropriate vector  
 15 containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SNEXN upon induction with isopropyl beta-D-  
 20 thiogalactopyranoside (IPTG). Expression of SNEXN in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SNEXN by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong  
 25 polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

30 In most expression systems, SNEXN is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham  
 35 Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from

WO 00/73334

PCT/US00/14831

SNEXN at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified SNEXN obtained by these methods can be used directly in the assay shown in Example XIV.

#### XI. Functional Assays

SNEXN function is assessed by expressing the sequences encoding SNEXN at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of SNEXN on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SNEXN and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding SNEXN and other genes of interest can be analyzed by

WO 00/73334

PCT/US00/14831

northern analysis or microarray techniques.

## **XII. Production of SNEXN Specific Antibodies**

SNEXN substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to  
 5 immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the SNEXN amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well  
 10 described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-  
 15 KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SNEXN activity by, for example, binding the peptide or SNEXN to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

## **XIII. Purification of Naturally Occurring SNEXN Using Specific Antibodies**

Naturally occurring or recombinant SNEXN is substantially purified by immunoaffinity  
 20 chromatography using antibodies specific for SNEXN. An immunoaffinity column is constructed by covalently coupling anti-SNEXN antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SNEXN are passed over the immunoaffinity column, and the column is  
 25 washed under conditions that allow the preferential absorbance of SNEXN (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SNEXN binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SNEXN is collected.

## **XIV. Identification of Molecules Which Interact with SNEXN**

SNEXN, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent.  
 30 (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SNEXN, washed, and any wells with labeled SNEXN complex are assayed. Data obtained using different concentrations of SNEXN are used to calculate values for the number, affinity, and association of  
 35 SNEXN with the candidate molecules.



Alternatively, molecules interacting with SNEXN are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

5 SNEYN may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	3	2124842	BRSTNOT07	817762T1 (OVRTUT01), 824400T1 (PROSNOT06), 2072065F6 (ISLTNOT01), 2124842H1 (BRSTNOT07), 2821317H1 (ADRETUT06), 3120257F6 (LUNGTUT13), 3703757F6 (PENCNOT07)
2	4	5215690	BRSTNOT35	508422R6 (TMLR3DT02), 568428R6 (MMLR3DT01), 568428T6 (MMLR3DT01), 1318278F1 (BLADNOT04), 1318278T1 (BLADNOT04), 1454211F1 (PENITUT01), 1741685R6 (HIPONON01), 2791135F6 (COLNTUT16), 2823783F6 (ADRETUT06), 5215690H1 (BRSTNOT35), SANA00022F1

WO 00/73334

PCT/US00/14831

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods
1	465	T36 S67 S174 S259 S383 T394 T414 T13 T77 S99 S160 S253 T291 S418 Y182 Y416		PhoX homology domain: I86-N205	Sorting nexin 14 [Homo sapiens] g4689266	BLAST-GenBank HMMER-Pfam MOTIFS
2	450	S108 T221 S43 S61 S63 S108 T190 S266 S288 S289 T69 T210 S337 T349 S354 T359 Y226 Y249 Y326	N192 N208	Signal peptide: M1-G37 PhoX homology domain: K56-E184	Sorting nexin 7 [Homo sapiens] g4689254	BLAST-GenBank SPScan HMMER-Pfam MOTIFS

WO 00/73334

PCT/US00/14831

Table 3

Polynucleotide SEQ ID NO:	Useful Fragments of Nucleic Acid Sequence	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
3	974-1018	Reproductive (0.191) Nervous (0.162) Gastrointestinal (0.147)	Cell Proliferation (0.574) Inflammation (0.338)	pINCY
4	165-209	Reproductive (0.200) Nervous (0.169) Gastrointestinal (0.154)	Cancer (0.569) Inflammation (0.169) Cell Proliferation (0.154)	pINCY

WO 00/73334

PCT/US00/14831

Table 4

Polynucleotide SEQ ID NO:	Library	Library Comment
3	BRSTNOT07	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
4	BRSTNOT35	Library was constructed using RNA isolated from breast tissue removed from a 46-year-old Caucasian female during a bilateral reduction mastoplasmy. Pathology indicated normal breast parenchyma, bilaterally. The patient presented with hypertrophy of breast and headache. Patient history included obesity, lumbago, glaucoma, and alcohol abuse. Family history included cataract, osteoarthritis, uterine cancer, benign hypertension, hyperlipidemia, alcoholic cirrhosis of the liver, cerebrovascular disease, and type II diabetes.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	PE Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	PE Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	PE Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score $\geq$ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 00/73334

PCT/US00/14831

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
  - a) an amino acid sequence of SEQ ID NO:1,
  - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence of SEQ ID NO:1,
  - c) a biologically active fragment of an amino acid sequence of SEQ ID NO:1, and
  - d) an immunogenic fragment of an amino acid sequence of SEQ ID NO:1.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence of SEQ ID NO:1.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence of SEQ ID NO:3.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
  - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
  - b) recovering the polypeptide so expressed.
10. An isolated antibody which specifically binds to a polypeptide of claim 1.



WO 00/73334

PCT/US00/14831

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4,
- b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4,
- c) a polynucleotide sequence complementary to a),
- d) a polynucleotide sequence complementary to b), and
- e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A pharmaceutical composition of claim 16, wherein the polypeptide comprises an amino acid sequence of SEQ ID NO:1.

WO 00/73334

PCT/US00/14831

18. A method for treating a disease or condition associated with decreased expression of functional SNEXN, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.

5 19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

10 20. A pharmaceutical composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional SNEXN, comprising administering to a patient in need of such treatment a pharmaceutical  
15 composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- 20 b) detecting antagonist activity in the sample.

23. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional SNEXN, comprising administering to a patient in need of such treatment a pharmaceutical  
25 composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim  
30 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying  
35 a compound that specifically binds to the polypeptide of claim 1.

WO 00/73334

PCT/US00/14831

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

5        b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a  
10        compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

15        a) exposing a sample comprising the target polynucleotide to a compound, and  
b) detecting altered expression of the target polynucleotide.

28. An isolated polynucleotide comprising a polynucleotide sequence of SEQ ID NO:4.

20        29. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 28.

30. A cell transformed with a recombinant polynucleotide of claim 29.

25        31. A transgenic organism comprising a recombinant polynucleotide of claim 29.

32. A method for producing a polypeptide comprising an amino acid sequence of SEQ ID NO:2, the method comprising:

a) culturing the cell of claim 30 under conditions suitable for expression of the polypeptide,  
30        and

b) recovering the polypeptide so expressed.

33. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 28, the method  
35        comprising:



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number  
**WO 00/73334 A3**(51) International Patent Classification<sup>7</sup>: **C12N 15/12**,  
C07K 14/47, 16/18, C12Q 1/68, A61K 38/17CA 95118 (US). **AZIMZAI, Yalda** [US/US]; 2045 Rock  
Springs Drive, Hayward, CA 94545 (US).

(21) International Application Number: PCT/US00/14831

(74) Agents: **HAMLET-COX, Diana et al.**; Incyte Genomics,  
3160 Porter Drive, Palo Alto, CA 94304 (US).

(22) International Filing Date: 26 May 2000 (26.05.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/136,740 27 May 1999 (27.05.1999) US  
60/139,566 16 June 1999 (16.06.1999) US(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE,  
ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP,  
KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD,  
MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD,  
SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ,  
VN, YU, ZA, ZW.(63) Related by continuation (CON) or continuation-in-part  
(CIP) to earlier applications:US 60/136,740 (CIP)  
Filed on 27 May 1999 (27.05.1999)  
US 60/139,566 (CIP)  
Filed on 16 June 1999 (16.06.1999)(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,  
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).**Published:**

— With international search report.

(71) Applicant (*for all designated States except US*): **INCYTE  
GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94304 (US).(88) Date of publication of the international search report:  
5 April 2001

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **YUE, Henry**  
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).  
**TANG, Y., Tom** [CN/US]; 4230 Ranwick Court, San Jose,*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: HUMAN SORTING NEXINS

(57) Abstract: The invention provides human sorting nexins (SNEXN) and polynucleotides which identify and encode SNEXN. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of SNEXN.

WO 00/73334 A3

WO 00/73334

PCT/US00/14831

PF-0706 PCT

## SEQUENCE LISTING

&lt;110&gt; INCYTE GENOMICS, INC.

YUE, Henry

TANG, Y. Tom

AZIMZAI, Yalda

&lt;120&gt; SORTING NEXINS

&lt;130&gt; PF-0706 PCT

&lt;140&gt; To Be Assigned

&lt;141&gt; Herewith

&lt;150&gt; 60/136,740

&lt;151&gt; 1999-05-27

&lt;150&gt; 60/139,566

&lt;151&gt; 1999-06-16

&lt;160&gt; 4

&lt;170&gt; PERL Program

&lt;210&gt; 1

&lt;211&gt; 465

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2124842CD1

&lt;400&gt; 1

Met	Tyr	Leu	Ile	His	Phe	Cys	Leu	Ile	Phe	Arg	Asn	Thr	Gln	Lys
1				5					10					15
Arg	Gly	Glu	Ser	Phe	Gly	Ile	Ser	Arg	Ile	Gly	Ser	Lys	Ile	Lys
				20					25					30
Gly	Val	Phe	Lys	Ser	Thr	Thr	Met	Glu	Gly	Ala	Met	Leu	Pro	Asn
				35					40					45
Tyr	Gly	Val	Ala	Glu	Gly	Glu	Asp	Asp	Phe	Ile	Glu	Glu	Gly	Ile
				50					55					60
Val	Val	Met	Glu	Asp	Asp	Ser	Pro	Val	Glu	Ala	Val	Ser	Thr	Pro
				65					70					75
Asn	Thr	Pro	Arg	Asn	Leu	Ala	Ala	Trp	Lys	Ile	Ser	Ile	Pro	Tyr
				80					85					90
Val	Asp	Phe	Phe	Glu	Asp	Pro	Ser	Ser	Glu	Arg	Lys	Glu	Lys	Lys
				95					100					105
Glu	Arg	Ile	Pro	Val	Phe	Cys	Ile	Asp	Val	Glu	Arg	Asn	Asp	Arg
				110					115					120
Arg	Ala	Val	Gly	His	Glu	Pro	Glu	His	Trp	Ser	Val	Tyr	Arg	Arg
				125					130					135
Tyr	Leu	Glu	Phe	Tyr	Val	Leu	Glu	Ser	Lys	Leu	Thr	Glu	Phe	His
				140					145					150
Gly	Ala	Phe	Pro	Asp	Ala	Gln	Leu	Pro	Ser	Lys	Arg	Ile	Ile	Gly
				155					160					165
Pro	Lys	Asn	Tyr	Glu	Phe	Leu	Lys	Ser	Lys	Arg	Glu	Glu	Phe	Gln
				170					175					180
Glu	Tyr	Leu	Gln	Lys	Leu	Leu	Gln	His	Pro	Glu	Leu	Ser	Asn	Ser
				185					190					195
Gln	Leu	Leu	Ala	Asp	Phe	Leu	Ser	Pro	Asn	Gly	Gly	Glu	Thr	Gln
				200					205					210

WO 00/73334

PCT/US00/14831

PF-0706 PCT

```

Phe Leu Asp Lys Ile Leu Pro Asp Val Asn Leu Gly Lys Ile Ile
215 220 225
Lys Ser Val Pro Gly Lys Leu Met Lys Glu Lys Gly Gln His Leu
230 235 240
Glu Pro Phe Ile Met Asn Phe Ile Asn Ser Cys Glu Ser Pro Lys
245 250 255
Pro Lys Pro Ser Arg Pro Glu Leu Thr Ile Leu Ser Pro Thr Ser
260 265 270
Glu Asn Asn Lys Lys Leu Phe Asn Asp Leu Phe Lys Asn Asn Ala
275 280 285
Asn Arg Ala Glu Asn Thr Glu Arg Lys Gln Asn Gln Asn Tyr Phe
290 295 300
Met Glu Val Met Thr Val Glu Gly Val Tyr Asp Tyr Leu Met Tyr
305 310 315
Val Gly Arg Val Val Phe Gln Val Pro Asp Trp Leu His His Leu
320 325 330
Leu Met Gly Thr Arg Ile Leu Phe Lys Asn Thr Leu Glu Met Tyr
335 340 345
Thr Asp Tyr Tyr Leu Gln Cys Lys Leu Glu Gln Leu Phe Gln Glu
350 355 360
His Arg Leu Val Ser Leu Ile Thr Leu Leu Arg Asp Ala Ile Phe
365 370 375
Cys Glu Asn Thr Glu Pro Arg Ser Leu Gln Asp Lys Gln Lys Gly
380 385 390
Ala Lys Gln Thr Phe Glu Glu Met Met Asn Tyr Ile Pro Asp Leu
395 400 405
Leu Val Lys Cys Ile Gly Glu Glu Thr Lys Tyr Glu Ser Ile Arg
410 415 420
Leu Leu Phe Asp Gly Leu Gln Gln Pro Val Leu Asn Lys Gln Leu
425 430 435
Thr Tyr Val Leu Leu Asp Ile Val Ile Gln Glu Leu Phe Pro Glu
440 445 450
Leu Asn Lys Val Gln Lys Glu Val Thr Ser Val Thr Ser Trp Met
455 460 465

```

<210> 2  
 <211> 450  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5215690CD1

```

<400> 2
Met Glu Gln Ala Pro Pro Asp Pro Glu Arg Gln Leu Gln Pro Ala
1 5 10 15
Pro Leu Glu Pro Leu Gly Ser Pro Asp Ala Gly Leu Gly Ala Ala
20 25 30
Val Gly Lys Glu Ala Glu Gly Ala Gly Glu Glu Ser Ser Gly Val
35 40 45
Asp Thr Met Thr His Asn Asn Phe Trp Leu Lys Lys Ile Glu Ile
50 55 60
Ser Val Ser Glu Ala Glu Lys Arg Thr Gly Arg Asn Ala Met Asn
65 70 75
Met Gln Glu Thr Tyr Thr Ala Tyr Leu Ile Glu Thr Arg Ser Val
80 85 90
Glu His Thr Asp Gly Gln Ser Val Leu Thr Asp Ser Leu Trp Arg
95 100 105
Arg Tyr Ser Glu Phe Glu Leu Leu Arg Ser Tyr Leu Leu Val Tyr
110 115 120

```

WO 00/73334

PCT/US00/14831

PF-0706 PCT

Tyr	Pro	His	Ile	Val	Val	Pro	Pro	Leu	Pro	Glu	Lys	Arg	Ala	Glu
				125					130					135
Phe	Val	Trp	His	Lys	Leu	Ser	Ala	Asp	Asn	Met	Asp	Pro	Asp	Phe
				140					145					150
Val	Glu	Arg	Arg	Arg	Ile	Gly	Leu	Glu	Asn	Phe	Leu	Leu	Arg	Ile
				155					160					165
Ala	Ser	His	Pro	Ile	Leu	Cys	Arg	Asp	Lys	Ile	Phe	Tyr	Leu	Phe
				170					175					180
Leu	Thr	Gln	Glu	Gly	Asn	Trp	Lys	Glu	Thr	Val	Asn	Glu	Thr	Gly
				185					190					195
Phe	Gln	Leu	Lys	Ala	Asp	Ser	Arg	Leu	Lys	Ala	Leu	Asn	Ala	Thr
				200					205					210
Phe	Arg	Val	Lys	Asn	Pro	Asp	Lys	Arg	Phe	Thr	Asp	Leu	Lys	His
				215					220					225
Tyr	Ser	Asp	Glu	Leu	Gln	Ser	Val	Ile	Ser	His	Leu	Leu	Arg	Val
				230					235					240
Arg	Ala	Arg	Val	Ala	Asp	Arg	Leu	Tyr	Gly	Val	Tyr	Lys	Val	His
				245					250					255
Gly	Asn	Tyr	Gly	Arg	Val	Phe	Ser	Glu	Trp	Ser	Ala	Ile	Glu	Lys
				260					265					270
Glu	Met	Gly	Asp	Gly	Leu	Gln	Ser	Ala	Gly	His	His	Met	Asp	Val
				275					280					285
Tyr	Ala	Ser	Ser	Ile	Asp	Asp	Ile	Leu	Glu	Asp	Glu	Glu	His	Tyr
				290					295					300
Ala	Asp	Gln	Leu	Lys	Glu	Tyr	Leu	Phe	Tyr	Ala	Glu	Ala	Leu	Arg
				305					310					315
Ala	Val	Cys	Arg	Lys	His	Glu	Leu	Met	Gln	Tyr	Asp	Leu	Glu	Met
				320					325					330
Ala	Ala	Gln	Asp	Leu	Ala	Ser	Lys	Lys	Gln	Gln	Cys	Glu	Glu	Leu
				335					340					345
Val	Thr	Gly	Thr	Val	Arg	Thr	Phe	Ser	Leu	Lys	Gly	Met	Thr	Thr
				350					355					360
Lys	Leu	Phe	Gly	Gln	Glu	Thr	Pro	Glu	Gln	Arg	Glu	Ala	Arg	Ile
				365					370					375
Lys	Val	Leu	Glu	Glu	Gln	Ile	Asn	Glu	Gly	Glu	Gln	Gln	Leu	Lys
				380					385					390
Ser	Lys	Asn	Leu	Glu	Gly	Arg	Glu	Phe	Val	Lys	Asn	Ala	Trp	Ala
				395					400					405
Asp	Ile	Glu	Arg	Phe	Lys	Glu	Gln	Lys	Asn	Arg	Asp	Leu	Lys	Glu
				410					415					420
Ala	Leu	Ile	Ser	Tyr	Ala	Val	Met	Gln	Ile	Ser	Met	Cys	Lys	Lys
				425					430					435
Gly	Ile	Gln	Val	Trp	Thr	Asn	Ala	Lys	Glu	Cys	Phe	Ser	Lys	Met
				440					445					450

&lt;210&gt; 3

&lt;211&gt; 1992

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2124842CB1

&lt;400&gt; 3

```

gtatgaaaac tccaaaagtt atggttagcta taccttttagt ttatcatttt caaaactggt 60
tttctttttt ttttaataatg tttttcaata aactagactg ttgtgataat ttgctatgta 120
cttgatacat ttttggtttta ttttcaggaa cacacagaaa aggggagaat catttggaat 180
cagcagaata ggttagcaaaa ttaaaggagt attcaaaagt accacaatgg agggagctat 240
gttgccctaat tatggtgtag ctgaagggtga agatgatttt attgaagaag gtattgttgt 300
aatggaagat gattctccag tggaggctgt gagcacacct aatactccc gaaaccttgc 360

```



WO 00/73334

PCT/US00/14831

PF-0706 PCT

```

tgcattgaaa attagcattc catatgtaga cttttttgag gatccctcct ctgaaaggaa 420
ggagaaaaaa gaaagaattc ctgtgttttg tattgatgtt gaaagaaatg atagaagagc 480
agttggacac gagcctgaac attggtctgt ctatagaaga tatcttgaat tctatgtact 540
tgaatcaaaa ctaacagaat ttcatgggtgc atttcttgat gccagcttc cttctaagag 600
gatcattggc cccaaaaatt atgaattctt aaagtcaaag aggggaagag tccaagaata 660
tctacagaaa cttctgcagc atccagaact gagtaatagt caacttctgg cagactttct 720
ttcccctaatt ggtggggaaa cacaatttct tgataagata ctaccagatg taaatcttgg 780
gaaaattata aaatctgttc ctggaaaact aatgaaagag aaaggtcagc atttggaacc 840
ttttatcatg aatttcatta attcttgtga gtctccaaag cctaaacca gtagaccaga 900
actgaccatt ctcagcccta cttcagaaaa caacaagaag cttttcaatg atctgtttaa 960
aaataatgca aaccgtgctg aaaatacaga gagaaagcaa aatcagaatt attttatgga 1020
ggtgatgact gtagaaggag tctatgatta cctgatgtat gtaggacggg tagttttcca 1080
ggttccctgac tggcttcac atctcttaat gggaaactcg atcctcttta aaaacaccct 1140
ggaaatgtat actgattact atcttcagtg taaactagaa cagctatttc aggagcaccg 1200
tttggctcct caataacac ttctcagaga tgctatatcc tgtgaaaaca ctgaacctcg 1260
ctctctccaa gaaacgaaa aaggagcaaa acagactttt gaagaaatga tgaattacat 1320
tccagatctg ttagtcaagt gtattggtga agaaaccaag tatgaaagca tcagacttct 1380
gtttgatggc ttacagcaac cagtactcaa caagcagctg acttatgttt tattggacat 1440
tgtgatacag gaactgtttc cagagctcaa aagggtacaa aagggaagtt cctctgtgac 1500
atcttggatg taacactttg gatttgggtat agaataaccc attgaaattt ctgctgtgag 1560
aggggtggtg aaatttactt ttttgggtat attcttatat atattatgta catcgctgtc 1620
tgaaatttta gttatttttt gtttttaata aagactaaca caaacttaat gattaaaagt 1680
gattgagtct catagtcttt catttgctag ctgtgatcca aattttatta gaacataagt 1740
cacttggatg tgccattttt attcataatg atgttatggc aaacagataa 1800
gactgataaa cttcgtattg tatagctttg aaaataatta tgcctagtat ggagaaacag 1860
gaataagatc tgattttctt agagttaata tatttttagt gattggtttt cttttttttt 1920
attttgtaca tagttaactg tgtatctata aataaagcat cctatatgag tttttaataa 1980
taaaaaaaaa aa 1992

```

&lt;210&gt; 4

&lt;211&gt; 2507

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5215690CB1

&lt;400&gt; 4

```

cgccgcgaac tgcagccatg gagcaggcac ctccggacc cgagcggcag ctccagccgg 60
cgcccttggg gccgctgggc tccccagac ctgggctggg ggctgcggtc ggcaaggaag 120
cggagggggc cggagaagag agctctgggg tcgacacgat gacacacaat aatttttgg 180
tgaagaagat agaaatcagt gtttcagaag cagaaaaacg aactggaaga aatgccatga 240
acatgcaaga aacatatact gcttacctca ttgaaacaag gtcagttgaa cataccgatg 300
gtcagagtgt cctaacagac tcaactatgg ggcgatatag tgaatttgag ttgttgagaa 360
gtacaccttt agtttactat ccacatattg ttgtgccacc tctgccagaa aaacgggcag 420
aatttggttg gcataaaact ctcttgagga ttgcttcaca tccatcctt tgtagagaca 540
ggattggttt agaaaaactt acacaggaag gtaactggaa ggagactgtg aatgaaactg 600
aaatcttcta tctgttttta tccagggtta aagcgcttaa tgcaacattc agagtgaata 660
ggtttcagct gaaggcagac gaccttaagc actatagtga tgaactgcag tctgtcatct 720
acccagacaa gagatttact gctagagtag cagatcgact ctatggtgta tataaagtac 780
cacatcttct tcgagtcaga gctagagtag ggagtgccat agaaaaagaa atgggtgatg 840
atgggaatta tggctcagttt ttcagtgaat tgtatgcac ttctattgat gatattttgg 900
gactgcagag catatgggat gatcagttaa aagagtatct tttttatgca gaagcattgc 960
aagatgaaga acattatgca gaacttatgc agtatgactt ggagatggct gctcaggact 1020
gggctgtgtg caggaacatc tgtgaggaac tggttaactg gactgtgaga acattctct 1080
tagcatccaa gaagcagcag ctctttggtc aagaaaactc agagcagaga gaagccagaa 1140
tgaaggggat agaagaacaa ataaatgaag gagaacaaca gctaaagtct aaaaatctgg 1200
taaagggtgc aacgcagagg ctcataagct atgcagtcac gcagatcagt atgtgcaaaa 1260
aaggcagaga atttgtgaaa aacgcagagg ctcataagct atgcagtcac gcagatcagt atgtgcaaaa 1320

```

WO 00/73334

PCT/US00/14831

PF-0706 PCT

```

agggaattca agtttggacc aatgctaagg aatgctttag caagatgtaa tcctgtgaat 1380
tgaattttct ttcaatcaaa gtgccccaaa acagaagcac aagtaaataa aagaaattta 1440
agtcactacc tagtatacat aaacatatac aataagttaa ataaattcag ctttctttta 1500
cttaattgtg gtcgtgttaa tgtagcacia aaaatatatt ttaatgaaga ttaaataatta 1560
taatttgagg ttttggggac tgggtgctgat tccaaaaagt taatttaata atatatacca 1620
acagattgtt tgtcacgctt ctgaaccaat gactgaatgt caagatgttc gtttaattct 1680
agatgtttgt ttcaagacca gctgtttcag atctattaat gtagggaatt tttccctaag 1740
attgaattcc tatatttact tggtaagatc cactaatctg ttatagggag ttgttttctc 1800
ttgccttata gttgagctat ttggtttgac aaagctcagc agaaacttga tgtgaaaaat 1860
ctacgatttt cttctacatt cattgatgcc ctttgtaatg tttgcataca ctgaaaatg 1920
ccttcaattt gtgttttacc agaattttga tactggtcag aaattttata ctgccaacaa 1980
agaagactca acttctcaga tctataatgg gatacattgt catcctctag caactcctat 2040
atagaaagtt ttaactgaat atgttacata taagaattaa attcttctca aataattctt 2100
aacctcagta atgagcctaa atttactctg cttggctctc tacacatggc atttcagggg 2160
ataagatgta gcatttcaat gtgtaagata tatgtactaa acatatgtgt tgctatcttc 2220
atcattaaca tccttctttt ctattgcttg gctgtaattt ttgtaaagat aaattatatt 2280
gtttttttgt atgtgtgttt gtagtatatg ttcagaaggc aagcatcttc attttgctag 2340
ctttgcagaa tcttaaaatg tgtactcgtt atttctaata atgtaaaaaa aaatccctag 2400
tcctgtttag catttgactt ttttatatgt tttaaatgtt gctggatttt tgtgctgttt 2460
gccaaactta tacaataaat aaatgaaata ttgtgctgaa aaaaaaa 2507

```

HPS Trailer Page  
for

# **Walk-Up\_Printing**

UserID: PML

Printer: cp2\_3d16\_gbkpptr

## **Summary**

<b>Document</b>	<b>Pages</b>	<b>Printed</b>	<b>Missed</b>
WO000073334	84	84	0
Total (1)	84	84	0

**Docket No.: PF-0706 USN**

## DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

## INTRACELLULAR SIGNALING MOLECULES

the specification of which:

/     / is attached hereto.

/ X / was filed on January 18, 2001 as application Serial No. 09/744,313 and if this box contains  
an X / \_\_/, was amended on \_\_\_\_\_.

/X / was filed as Patent Cooperation Treaty international application No. PCT/US00/14831 on May 26, 2001, if this box contains an X / /, was amended on under Patent Cooperation Treaty Article 19 on \_\_\_\_\_ 2001, and if this box contains an X / /, was amended on \_\_\_\_\_.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

**Docket No.: PF-0706 USN**

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	// Yes // No
_____	_____	_____	// Yes // No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/136,740	27 May 1999	Expired
60/139,566	16 June 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
_____	_____	_____

I hereby appoint the following:

12 Lucy J. Billings  
Michael C. Cerrone  
Diana Hamlet-Cox  
Richard C. Ekstrom  
Barrie D. Greene  
Lynn E. Murry  
Shirley A. Recipon  
Susan K. Sather  
Michelle M. Stempien  
David G. Streeter  
Stephen Todd  
P. Ben Wang

Reg. No. 36,749  
Reg. No. 39,132  
Reg. No. 33,302  
Reg. No. 37,027  
Reg. No. 46,740  
Reg. No. 42,918  
Reg. No. 47,016  
Reg. No. 44,316  
Reg. No. 41,327  
Reg. No. 43,168  
Reg. No. 47,139  
Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0706 USN

**LEGAL DEPARTMENT**  
**INCYTE GENOMICS, INC.**  
**3160 PORTER DRIVE, PALO ALTO, CA 94304**

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

**Sole Inventor or  
First Joint Inventor:**

*rw*

Full name:

HENRY YUE

Signature:

*Henry Yue*

Date:

January 26, 2001

Citizenship

United States of America

Residence:

Sunnyvale, California

CA

P.O. Address:

826 Lois Avenue  
Sunnyvale, California 94087

**Second Joint Inventor:**

*2w*

Full name:

Y. TOM TANG

Signature:

*Y. Tom Tang*

Date:

February 27, 2001

Citizenship

People's Republic of China USA

Residence:

San Jose, California

CA *U. Tang*

P.O. Address:

4230 Ranwick Court  
San Jose, California 95118

Docket No.: PF-0706 USN

Third Joint Inventor:

ZW

Full name:

YALDA AZIMZAI

Signature:

Yalda Azimzai

Date:

February 2nd, 2001

Citizenship

United States of America

Residence:

Castro Valley, California CA

P.O. Address:

5518 Boulder Canyon Drive  
Castro Valley, California 94552